Pretrained Convolutional Neural Network for Fruit Classification Analysis of Pineapple Plantation Images

Nurhazirah Mohd Rahim
College of Engineering, Universiti Teknologi MARA, Johor Branch, 81750 Masai, Johor, Malaysia
2024581393@student.uitm.edu.my

Muhammad Asraf Hairuddin
College of Engineering, Universiti Teknologi MARA, Johor Branch, 81750 Masai, Johor, Malaysia | Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
masraf@uitm.edu.my (corresponding author)

Megat Syahirul Amin Megat Ali
College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
megatsyahirul@uitm.edu.my

Nooritawati Md. Tahir
School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia | Department of Cybersecurity, International Information Technology University, Almaty, Kazakhstan
noori425@uitm.edu.my

Ali Abd Almisreb
Department of Computer Science and Engineering, Faculty of Engineering and Natural Sciences, International University of Sarajevo, Sarajevo 71210, Bosnia and Herzegovina
alimes96@yahoo.com

Nur Dalila Khirul Ashar
College of Engineering, Universiti Teknologi MARA, Johor Branch, 81750 Masai, Johor, Malaysia | Department of Computer and Communication Systems, Faculty of Engineering, Universiti Putra Malaysia, Malaysia
nurdalila306@uitm.edu.my

Received: 12 October 2024 | Revised: 22 December 2024 | Accepted: 29 December 2024

 DOI: https://doi.org/10.48084/etasr.9249

ABSTRACT

The adoption of precision agriculture in pineapple farming has a significant impact by increasing the yield and reducing the input resources while improving the management of pineapple crops. The intersection of advanced drone technology and cutting-edge artificial intelligence has reformed fruit crop management through revolutionary levels of automation, precision fruit detection, yield estimation, and crop health detection. However, the capability for obscuring the detection of subtle features to better manage occlusions and complex environments in images captured by drones at certain heights with drones is challenging to distinguish, thus hindering an accurate object analysis for fruit-environment differentiation. The proposed work uses Deep Learning (DL) techniques to classify pineapple fruit images captured ten meters above the ground. This is achieved specifically through the use of pretrained models and Faster Region-Based Convolutional Neural Networks (Faster R-CNNs) due to their ability to learn robust

interpretations from images for object classification tasks. This paper evaluates the capabilities and accuracies of four pretrained models, namely ResNet-101, ResNet-50, Inception-ResNet-v2, and VGG-19, to detect and classify the pineapple fruit amidst the complex background and varying lighting conditions. By evaluating the pretrained models for pineapple fruit classification using comprehensive metrics (True Positive Rate (TPR), False Positive Rate (FPR), Accuracy (ACC), Recall (REC), Precision (PRE), F1-score), the results reveal that the Faster R-CNN architecture with the VGG-19 pretrained model outperformed the other architectures, demonstrating the best performance in pineapple fruit detection with an ACC of 0.7924 (79.24%), a PRE of 0.9990 (99.90%), a REC of 0.7930 (79.30%), and an F1-score of 0.8839 (88.39%). The effectiveness of this model in overseeing complex scenarios suggests potential improvements in classification accuracy compared to other pretrained models, while acknowledging performance variability across various architectures.

Keywords-ananas comosus; deep learning; image processing; pretrained deep neural networks; pineapple farming

I. INTRODUCTION

The Ananas comosus (or pineapple) fruit is a commercially grown tropical fruit in high demand in local and foreign markets, making its cultivation and production a key component of the economies of many countries. Notably, the growing interest in pineapple as a high-value crop has become a viable source of income for smallholder farmers [1, 2]. As technology advances, the scope of precision agriculture in pineapple production is likely to expand and bring greater benefits to the sustainability and productivity of the global pineapple industry. Thus, yield estimates can be monitored with detailed insights into the crop conditions and environmental factors, facilitating timely interventions and informed crop management decisions [3]. Traditional methods of estimating pineapple crop yields rely heavily on manual processes that are inherently subjective and generally prone to yield prediction bias. In most cases, agricultural officials use visual techniques that require a great deal of human judgment and experience. Such subjectivity introduces large yield errors that influence agricultural management decisions and productivity. Therefore, it is necessary to increase the efficiency of pineapple farming. To date, several approaches are available to automate current farming practices, namely remote sensing [4-6], computer vision [7-9], and Machine Learning (ML) algorithms [10-12].

Recent studies incorporating automated approaches and image analysis have been actively conducted to improve the accuracy and efficiency of pineapple crop classification processes. Authors in [13] proposed an enhanced segmentation approach that integrates image thresholding and Hue Value Segmentation (HVS) color space transformation for crop counting in pineapple plantations. This effectively identifies features in images with low contrast, which may result from varying heights and lighting conditions in the captured color image. Meanwhile, authors in [14] developed a recognition algorithm using the YOLO-v4 model to improve accuracy and reduce training time in predicting fruit maturity before harvesting. However, the proposed approach requires a massive dataset to obtain a robust model. Authors in [15] proposed a combined method consisting of Cascade Object Detector (COD), HVS, Adaptive Red and Blue Chromatic Map (ARB), Normalized Difference Index (NDI), and Convolutional Neural Network (CNN). The result is a high-quality image for accurate maturity identification and segmentation of pineapples from the background. At the same time, authors in [16] proposed counting and localizing flowering pineapple plants automatically by analyzing the density of flowers in the images using a U-net backbone model. This results in accurate counting performance with low error, which improves the efficiency and effectiveness of pineapple harvesting operations.

Precision agriculture, as a part of advanced agricultural techniques utilizing drone image processing and Deep Learning (DL) models, is one of the latest approaches currently being explored to manage agronomic variability for environmental sustainability, improved crop quality, and efficient management practices [17-19]. With regard to pineapple farming, DL techniques can be employed to classify the pineapple characteristics. This includes maturity and quality based on color, shape, and texture [20], detection of pineapple fruits and flowers in cluttered backgrounds [21], and accurate automated detection of ripe pineapples and their three-dimensional (3D) distance in fields [22]. In [23], a DL method, which is a layered network architecture based on CNNs, was employed to classify 20 unique types of fruits and vegetables using deep CNNs with data augmentation techniques. In addition, authors in [24] demonstrated a detailed performance evaluation of several CNN models for fruit and vegetable detection, which showed significant improvements with modern DL approaches in terms of both accuracy and speed. These results substantiate that CNNs are essential in achieving classification performance in advanced levels of agriculture. Remarkably, many studies apply CNNs to this field, demonstrating the outstanding performance and practical results of CNNs in agricultural technology.

Despite the cutting-edge capabilities available in image processing, there is still room for improvement as the quality of captured images deteriorates. This includes blurring with loss of detail due to low resolution images produced during a heightened flight, which affects model performance. In particular, pretrained CNNs have attracted research interest in crop classification due to their functionality and effectiveness in producing an accurate classification based on visual characteristics captured in images. The model has demonstrated superior capabilities in fruit classification tasks for various crops. However, the performance of pretrained models requires trial and error, as no specific model fits all crops. Furthermore, variations in crop appearance and features, influenced by environmental conditions, differences in dataset size and quality, and specific classification requirements, can significantly affect the performance of pretrained models. For example, authors in [25] discovered that VGG-19 outperformed InceptionV3 and MobileNetV2 in ripeness classification of the Philippine guyabano fruit, achieving an impressive accuracy of 99.25%. In addition, authors in [26] demonstrated how various pretrained networks performed in pineapple classification. The VGG19 model was found to be excellent at identifying the finer details and patterns in dragon fruit images. Similarly, authors in [27] evaluated ResNet-50 and VGG-19 for fruit quality classification, with VGG-19 yielding the best performance with over 94% accuracy. Authors in [28] presented the result of ResNet-50-v2, which achieved 98.89% accuracy on a dataset of 41 fruit categories. The pretrained CNN has demonstrated its effectiveness in classifying fused images, with ResNet-50 demonstrating superior performance in both accuracy and convergence speed [29]. Moreover, the VGG-19 model excels at detecting finer details and patterns in images, especially in the agricultural field, for classifying fruits and vegetables against crowded backgrounds. For instance, authors in [30] highlighted that VGG-19 can detect fine details that may be obscured by the surrounding elements, and then effectively classify images, even when the latter contain overlapping foliage or other distractions. The VGG19 could effectively segment and recognize objects in images that present challenges due to varying colors and lighting conditions [31, 32].

This research assesses how effectively the pretrained Faster R-CNNs classify annotated images of pineapple fruits in the presence of complex backgrounds and varying lighting conditions. Accordingly, comparative evaluations between ResNet-101, ResNet-50, Inception-ResNet-v2, and VGG-19 provide valuable information on how the models perform with different pretrained networks for pineapple classification. Performance metrics, such as TPR, FPR, ACC, REC, PRE, and F1-score are used to assess the model performance in an analysis of CNN for image classification. These metrics provide valuable insights into the model's capability to correctly identify the positives and negatives of pineapple detection, as well as its overall effectiveness in classification tasks.

II. METHODOLOGY

The present study proposes the analysis of pineapple image data collected by drones, such as the DJI Phantom 3 advanced quadcopter, which is used to capture high-quality images with a 4K resolution RGB camera [33]. The analysis of the drone images was conducted on a personal computer equipped with a 2.1 GHz Intel Core i5-13420H CPU, 6 GB GDDR6 RAM, NVIDIA GeForce RTX 4050, and MATLAB R2024a. The area where drone images were collected based on direct field observation is located in a pineapple plantation in Johor, Malaysia. Furthermore, a computational analysis was conducted on images captured from ten meters above the ground. It utilizes a two-stage algorithm that first generates region proposals through a Region Proposal Network (RPN) and subsequently classifies these samples using a Faster R-CNN. This approach effectively streamlines the object detection process by leveraging the strengths of the RPN and the Faster R-CNN framework, known for its efficiency in identifying pineapple fruits as objects within images.

The execution of the algorithm begins with image processing and the overall workflow is presented in Figure 1. The dataset is divided into training and testing sets. A pretrained model, previously trained on a large dataset, is employed to extract a feature map from the pineapple images, enabling it to effectively identify various image features. This approach leverages transfer learning, allowing the model to utilize its prior knowledge to improve its performance in the specific task of pineapple image classification. The feature map is processed by the RPN, which identifies potential Regions Of Interest (ROIs) in the image. In ROI pooling, the identified regions are pooled to ensure that they are of uniform size, which is necessary for further processing. In addition, the pooled regions are passed through a classifier to determine the class of objects within the regions (e.g., pineapple or non-pineapple). The model performance is evaluated on the testing set, which is necessary to validate the model and ensure its accuracy and reliability in classifying pineapples in plantation images. Moreover, the pineapple fruit counts obtained from the automated analysis of the image frames were manually verified by experts for comparison. This overcomes the limitations of the traditional assessments of the counting process and strengthens the reliability of the results.

A. Dataset Preparation

Two datasets are required, commonly designated as training and testing images. During the training phase, the Faster R-CNN learns these training images to recognize patterns and features that distinguish the pineapples from the background. In contrast, the testing images are datasets used to evaluate the performance of the trained model. The latter is separate from the training set and inaccessible to the model during the training phase. Meanwhile, the testing images are utilized to assess how well the Faster R-CNN generalizes to new data. It should be noted that the proposed model performs classification on the test images, and these outcomes are compared with the true labels using performance metrics. For this study, a total of 300 images of pineapple fruits, each with a size of 240×300 pixels and various positions and shapes, were selected as the training images. All the training images were annotated as pineapple using the Make Sense software. For the testing images, an additional set of ten frames containing multiple pineapples captured from a height of ten meters above the ground was selected and evaluated.

B. Faster Region-based Convolutional Neural Networks

The pineapple images were trained with the pretrained models, namely ResNet-101, ResNet-50, Inception-ResNet-v2, and VGG-19 as pretrained Deep Neural Networks (DNNs) to save time and computational resources compared to training a Faster R-CNN from scratch, which is especially important when working with limited datasets or when the computational resources are limited. The classifier uses an anchor box mechanism to oversee numerous sizes and aspect ratios.

1) ResNet-101 and ResNet-50

ResNet-101 and ResNet-50 are variations of the Residual Network (ResNet) architecture designed to enable effective CNN training. Specifically, the ResNet-101 consists of 101 layers, while the ResNet-50 contains only 50 layers. Both architectures use residual connections, which help mitigate the vanishing gradient problem and enable the effective training of deeper networks. Moreover, the ResNet-101 contains a deeper architecture, which allows it to capture more intricate features and perform slightly better on complex visual or image processing tasks. In contrast, the ResNet-50 is more computationally efficient and has a lower processing time.

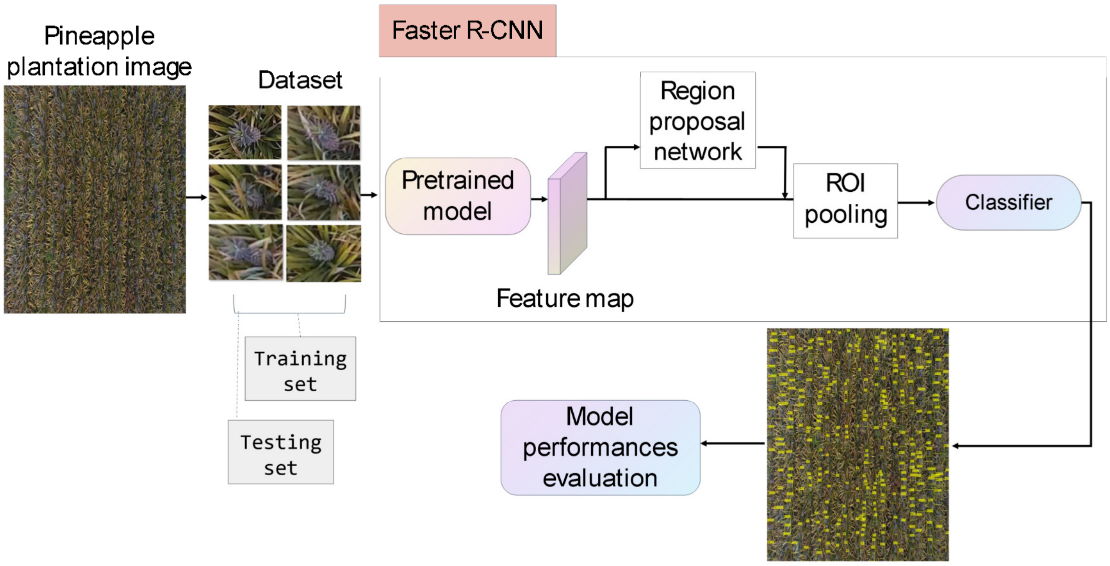


Fig. 1. Workflow of the proposed study.

2) Inception-ResNet-v2

Inception-ResNet-v2 incorporates a CNN architecture that combines the multi-scale feature extraction of the inception module. This is done by utilizing inception modules with varying filter sizes within each layer and integrating them with the residual connections of ResNet to facilitate the gradient flow in deep networks. In addition, this architecture hybridizes individual components in order to enhance the model accuracy, adaptability, and ultimately the computational efficiency in image classification tasks.

3) VGG-19

VGG-19 is a 19-layer CNN architecture from the Visual Geometry Group (VGG), consisting of 16 convolutional layers and three fully connected layers. This architecture is based on a simple and homogeneous structure, where each convolutional layer has a filter size of 3 × 3 and is followed by a max-pooling layer to decrease the spatial dimensions. This architecture is suitable for image classification tasks due to its deep layering and relatively small convolution filters, which assist in capturing detailed image patterns.

C. Model Performance Evaluation

Several performance indicators are used to examine and evaluate the performance of Faster R-CNN by examining the TPR, FPR, ACC, PRE, and F1-score, shown in (1)-(5). True Positive (TP) can be defined as correctly predicted as true pineapple, True Negative (TN) represents the correctly predicted as non-pineapple, False Positive (FP) represents incorrectly predicted as pineapple, and False Negative (FN) represents incorrectly predicted as non-pineapple. Based on these parameters, the performance metrics are then calculated to benchmark which pretrained DNN has the best performance in performing classification for the pineapples within the plantation.

$$\mathrm{TPR\,/\,REC} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{1}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} \tag{2}$$

$$\mathrm{ACC} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{3}$$

$$\mathrm{PRE} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{4}$$

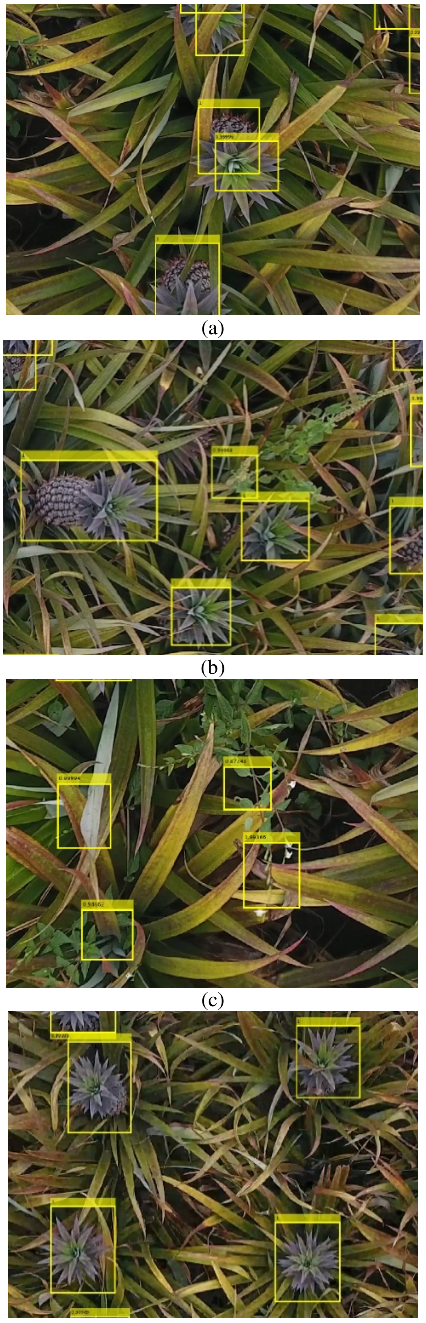
$$\mathrm{F1-score} = 2\times\frac{\mathrm{PRE}\times\mathrm{REC}}{\mathrm{PRE}+\mathrm{REC}} \tag{5}$$

III. RESULTS AND DISCUSSION

The results for all the pretrained DNNs are exhibited in Figure 2. As displayed in Figure 2(c), with Inception-ResNet-v2, it can be observed that all the bounding boxes in the sample frame misclassified the leaves or background as pineapples, reflecting poor model accuracy. However, in Figures 2(a), 2(b) and 2(d), most of the bounding boxes were correctly placed on the pineapple fruit in the image, demonstrating a better classification rate. It can be observed that the proper selection of pretrained DNN models plays an important role in the successful recognition of pineapple fruits. Table I compares the computational time required for each model to process the training and testing images.

TABLE I. TRAINING AND TESTING TIMES

Pretrained model	Training time (s)	Testing time (s)
ResNet-101	2880.6	1.30
ResNet-50	2842.2	1.04
Inception-ResNet-v2	6195.6	1.84
VGG-19	678.6	5.64



(a)

(b)

(c)

(d)

Fig. 2. Demonstration of pineapple detection from the sample frame using the pretrained models: (a) ResNet-101, (b) ResNet-50, (c) Inception-ResNet-v2, and (d) VGG-19.

The execution testing time considers the average time to process ten frames of cropped pineapple images. The process of fine-tuning the model on this new dataset of testing images involves adapting the learned features from the training phase to the specific characteristics through iterative training. The longest training time is for Inception-ResNet-v2, whereas the longest testing time is for VGG-19. The reason for this is probably that Inception-ResNet-v2 encompasses a complex architecture designed to extract intricate features, which requires more computations during training to adjust for a substantial number of parameters. VGG-19 consists of deep stacks of convolutional layers utilizing small receptive fields (3 × 3 filters). This results in a substantial number of operations during inferencing, and therefore requires significant memory and computational resources to process each test image. Although Inception-ResNet-v2 takes the longest to train, the classification performance is poor when running on test images. Thus, the longer training time is no guarantee of good quality results.

Table II provides the recorded values for the performance metrics, including the number of fruits manually counted (n), TP, TN, FP, FN, ACC, PRE, REC, and F1-score. The highest ACC signifies that nearly all predictions regarding the detection and classification of the object (pineapples) are correct. Similarly, the highest PRE indicates a minimum number of FPs, whereas the highest REC reflects a minimum number of FNs. A top F1-score demonstrates an optimal balance between PRE and REC, suggesting that the model is both reliable and suitable for real-world applications, where both metrics are crucial. This suggests that the model effectively identifies positive cases (high REC) while generating few false alarms (high PRE). Therefore, a high F1-score is indicative of a robust overall performance. The results presented in Table II compare the performance of four pretrained models, namely ResNet-101, ResNet-50, Inception-ResNet-v2, and VGG-19, on a test image across ten different frames using the performance metrics. It is noteworthy that ResNet-101 exhibits significantly varied performance across frames. Frame 1 stands out with exceptional results, achieving an ACC, PRE, REC, and F1-score of 0.9544, 0.9941, 0.9599, and 0.9767, respectively. However, the subsequent frames show a significantly lower performance, with ACC ranging from 0.5876 to 0.6630. REC remains consistently high at nearly 1.0, indicating that the model captures almost all positive instances. However, it struggles with PRE in later frames. Meanwhile, ResNet-50 demonstrates more consistent performance compared to ResNet-101. Its ACCs range from 0.7088 to 0.7721, which is significantly more stable across frames. Equivalently to ResNet-101, it maintains a perfect REC of 1.0 in most frames, suggesting robust positive instance detection. Frame 1 presents the highest ACC and REC at 0.7721, while other frames maintain ACC above 0.7, indicating more reliable performance. The Inception-ResNet-v2 model exhibits the most challenging overall detection performance among the four pretrained models. In particular, the ACCs are significantly lower, ranging from 0.4108 to 0.4831. REC varies between 0.4419 and 0.4971, indicating consistent detection capabilities. Nevertheless, it remains relatively low compared to PRE, which varies between 0.5390 and 0.9448. The F1-scores are mostly between 0.5824 and 0.6514. Frame 10 appears to be the best performing frame for this architecture. At the same time, VGG-19 emerges as the second-best performing model after the first frame of ResNet-101. It demonstrates consistently high ACCs ranging from 0.7493 to 0.8402. Frame 7 is particularly impressive, achieving an ACC of 0.8402 and an F1-score of 0.9132. Similar to other models, the VGG-19 maintains near-perfect REC across frames. Considering all the metrics, VGG-19 provides robust and reliable performance across most frames. The results suggest that the model architecture has a significant impact on performance, and that performance can vary significantly across different frames or data subsets. Therefore, a careful evaluation across multiple frames or scenarios is crucial when selecting a model for a particular task. To better visualize the performance metrics, the average performance for ten frames of test images is depicted in Figure 3.

TABLE II. PERFORMANCE METRICS

Pretrained model	Test image	n	TP	TN	FP	FN	ACC	PRE	REC	FI-score
ResNet-101	Frame 1	351	335	0	2	14	0.9544	0.9941	0.9599	0.9767
	Frame 2	364	229	0	0	135	0.6291	1.0000	0.6291	0.7723
	Frame 3	370	225	0	1	144	0.6081	0.9956	0.6098	0.7563
	Frame 4	355	219	0	0	136	0.6169	1.0000	0.6169	0.7631
	Frame 5	362	240	0	1	121	0.6630	0.9959	0.6648	0.7973
	Frame 6	355	225	0	0	130	0.6338	1.0000	0.6338	0.7759
	Frame 7	363	223	0	2	138	0.6143	0.9911	0.6177	0.7611
	Frame 8	367	224	0	1	142	0.6104	0.9956	0.6120	0.7580
	Frame 9	363	216	0	1	146	0.5950	0.9954	0.5967	0.7461
	Frame 10	334	208	0	0	146	0.5876	1.0000	0.5876	0.7402
ResNet-50	Frame 1	351	271	0	0	80	0.7721	1.0000	0.7721	0.8714
	Frame 2	364	258	0	1	105	0.7088	0.9961	0.7107	0.8296
	Frame 3	370	264	0	0	106	0.7135	1.0000	0.7135	0.8328
	Frame 4	355	261	0	1	93	0.7352	0.9962	0.7373	0.8474
	Frame 5	362	272	0	0	90	0.7514	1.0000	0.7514	0.8580
	Frame 6	355	258	0	0	97	0.7268	1.0000	0.7268	0.8418
	Frame 7	363	260	0	1	102	0.7163	0.9962	0.7182	0.8347
	Frame 8	367	265	0	0	102	0.7221	1.0000	0.7221	0.8386
	Frame 9	363	262	0	0	101	0.7218	1.0000	0.7218	0.8384
	Frame 10	354	254	0	3	97	0.7175	0.9883	0.7236	0.8355
Inception-ResNet-v2	Frame 1	351	149	0	25	177	0.4245	0.8563	0.4571	0.5960
	Frame 2	364	157	0	20	187	0.4313	0.8870	0.4564	0.6027
	Frame 3	370	152	0	26	192	0.4108	0.5390	0.4419	0.5824
	Frame 4	355	163	0	16	176	0.4592	0.9106	0.4808	0.6293
	Frame 5	362	154	0	16	192	0.4254	0.9059	0.4451	0.5969
	Frame 6	355	158	0	21	176	0.4451	0.8827	0.4731	0.6160
	Frame 7	363	167	0	10	186	0.4601	0.9435	0.4731	0.6302
	Frame 8	367	158	0	13	196	0.4305	0.9240	0.4463	0.6019
	Frame 9	363	157	0	14	192	0.4325	0.9181	0.4499	0.6038
	Frame 10	354	171	0	10	173	0.4831	0.9448	0.4971	0.6514
VGG-19	Frame 1	351	266	0	0	85	0.7578	1.0000	0.7578	0.8622
	Frame 2	364	289	0	0	75	0.7940	1.0000	0.7940	0.8851
	Frame 3	370	287	0	1	82	0.7757	0.9965	0.7778	0.8737
	Frame 4	355	284	0	0	71	0.8000	1.0000	0.8000	0.8889
	Frame 5	362	282	0	0	80	0.7790	1.0000	0.7790	0.8758
	Frame 6	355	282	0	0	73	0.7944	1.0000	0.7944	0.8854
	Frame 7	363	305	0	1	57	0.8402	0.9967	0.8425	0.9132
	Frame 8	367	302	0	0	65	0.8229	1.0000	0.8229	0.9028
	Frame 9	363	272	0	1	90	0.7493	0.9963	0.7514	0.8567
	Frame 10	354	287	0	0	67	0.8107	1.0000	0.8107	0.8955

More comprehensive assessments comparing various architectures should be performed on diverse test images. Overall, the VGG-19 model achieved the highest performance across all metrics, with a mean ACC, mean PRE, mean REC, and mean F1-score of 0.7924, 0.9990, 0.7939, and 0.8839, respectively. This indicates that VGG-19 was able to correctly classify most of the test images while minimizing FP and FN. ResNet-50 also performed well, with a mean ACC, mean PRE, mean REC, and mean F1-score of 0.7285, 0.9977, 0.7297, and 0.8428, respectively. This suggests that ResNet-50 is a strong contender for the pineapple image classification task. These metrics still demonstrate the effectiveness of ResNet-50 in accurately classifying the test samples and maintaining a good balance between PRE and REC, although slightly less accurate than VGG-19. ResNet-101, despite having a deeper architecture than ResNet-50, had lower performance across all metrics. Its mean ACC, mean PRE, mean REC, and mean F1-score were 0.6513, 0.9968, 0.6528, and 0.7847, respectively. This could be due to overfitting the training data or the specific characteristics of the test image. Inception-ResNet-v2 had the lowest performance among the four models, with a mean ACC, mean PRE, mean REC, and mean F1-score of 0.4402, 0.9027, 0.4621, and 0.6111, respectively. Despite the high mean PRE and low mean REC, these metrics correctly classify less than half of the test samples, indicating that they fail to identify a significant portion of the positive samples. This suggests that the Inception-ResNet-v2 architecture may not be well suited to this classification task or the test image.

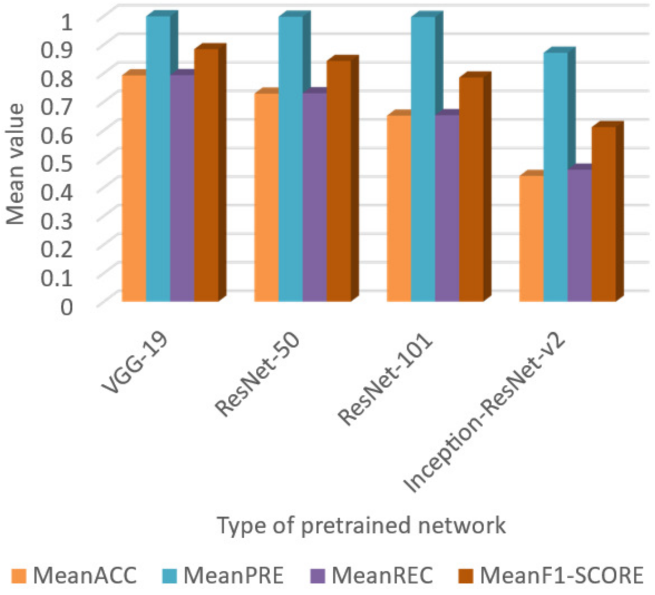


Fig. 3. Average performance comparison across multiple metrics.

IV. CONCLUSION

In conclusion, based on the provided performance metrics, the proposed detection model uses several pretrained learning models, namely ResNet-50, ResNet-101, Inception-ResNet-v2, and VGG-19 in which building a highly generalizable model to examine various types of objects can improve the classification outcome and further expand to perform detection tasks. It should be noted that the main challenge that needs to be addressed during the fruit image classification process is the poor quality of the images captured by the drone. Their quality can degrade, resulting in blurriness, loss of morphological details, and low resolution because the image is captured during a heightened flight. To overcome this, the model can be refined to ensure that the image classification system is less sensitive to the loss of image details. Hence, the proposed method is applied to the problem of classifying pineapple fruits in ten frames of images taken by a drone at ten meters above the ground. The performance metrics indicate that the model reliably identifies the pineapple fruit in images with complex backgrounds and varying lighting conditions, while highlighting the importance of selecting an appropriate pretrained model for a given computer vision task and dataset. The present study highlighted that VGG-19 as a backbone network performed best at extracting features from images and recognizing objects in complex environments, which enables the identification of an object in a complex background. In addition, the VGG-19 architecture combined with a Faster Region-Based Convolutional Neural Network (Faster R-CNN) was shown to improve the Accuracy (ACC) and robustness of detection and classification, making it as the preferred option for applications that require high Precision (PRE) in object detection in the presence of complex backgrounds. Moreover, a high F1-score indicates that the model is reliable and suitable for real-world applications where both aspects are critical, as it accurately identifies positive cases with high Recall (REC) and generates infrequent false alarms with high PRE. Thus, a high mean F1-score of 0.8839 or 88.39% reflects a robust overall performance. In contrast, ResNet-50. ResNet-101, and Inception-ResNet-v2 demonstrate lower performance, with Inception-ResNet-v2 having the weakest results. Therefore, these models are not a good choice for object classification under complex background conditions. In future work, a more thorough analysis considering diverse datasets when evaluating the models would provide a more comprehensive assessment and a generalized model. Furthermore, improved models should be accompanied by explainability techniques to understand why models produce certain performance outcomes when identifying regions, leading to targeted improvements.

ACKNOWLEDGMENT

This research work was funded by the Ministry of Higher Education (MOHE), Malaysia under the Fundamental Research Grant Scheme FRGS/1/2023/TK08/UITM/02/10. The authors also acknowledge the support from Universiti Teknologi MARA.

REFERENCES

[1] W. A. I. U. Bandara, K. A. D. K. S. Kuruppuarachchi, N. N. D. Maduwantha, U. S. S. Samaratunge Arachchillage, T. A. D. T. N. D. Alwis, and T. A. Kuruppu, "Smart Intelligent Pineapple Farming Assistant Agent (SIPFAA)," in *2022 4th International Conference on Advancements in Computing*, Colombo, Sri Lanka, 2022, pp. 48–53, https://doi.org/10.1109/ICAC57685.2022.10025103.

[2] N. H. M. Suhaimi and F. A. Fatah, "Profitability of Pineapple Production (Ananas comosus) among Smallholders in Malaysia," *International Journal of Recent Technology and Engineering*, vol. 8, no. 4, pp. 4201–4207, Nov. 2019, https://doi.org/10.35940/ijrte.D7780.118419.

[3] F. Alaieri, "Precision Agriculture based on Machine Learning and Remote Sensing Techniques," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14206–14211, Jun. 2024, https://doi.org/10.48084/etasr.6986.

[4] M. Gonzalez-Hernandez *et al.*, "Evaluation of the Influence of Multispectral Imaging for Object Detection in Pineapple Crops," in *2023 IEEE 5th International Conference on BioInspired Processing*, San Carlos, Alajuela, Costa Rica, 2023, pp. 1–6, https://doi.org/10.1109/BIP60195.2023.10379335.

[5] A. Sakuma and H. Yamano, "Satellite Constellation Reveals Crop Growth Patterns and Improves Mapping Accuracy of Cropping Practices for Subtropical Small-Scale Fields in Japan," *Remote Sensing*, vol. 12, no. 15, Aug. 2020, Art. no. 2419, https://doi.org/10.3390/rs12152419.

[6] T. D. Pham, D. T. H. Nguyen, D. H. Tran, and T. T. Tran, "Application of satellite images and GIS to assess the site situation and propose solutions for developing pineapple cultivation in U Minh Thuong, Kien Giang, Vietnam," *IOP Conference Series: Earth and Environmental Science*, vol. 1306, no. 1, Mar. 2024, Art. no. 012026, https://doi.org/10.1088/1755-1315/1306/1/012026.

[7] F. Meng, J. Li, Y. Zhang, S. Qi, and Y. Tang, "Transforming unmanned pineapple picking with spatio-temporal convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 214, Nov. 2023, Art. no. 108298, https://doi.org/10.1016/j.compag.2023.108298.

[8] M. Mohd Ali, N. Hashim, S. K. Bejo, M. Jahari, and N. A. Shahabudin, "Innovative non-destructive technologies for quality monitoring of pineapples: Recent advances and applications," *Trends in Food Science & Technology*, vol. 133, pp. 176–188, Mar. 2023, https://doi.org/10.1016/j.tifs.2023.02.005.

[9] S. Alqethami, B. Almtanni, W. Alzhrani, and M. Alghamdi, "Disease Detection in Apple Leaves Using Image Processing Techniques," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8335–8341, Apr. 2022, https://doi.org/10.48084/etasr.4721.

[10] G. Qiu *et al.*, "Nondestructive Detecting Maturity of Pineapples Based on Visible and Near-Infrared Transmittance Spectroscopy Coupled with Machine Learning Methodologies," *Horticulturae*, vol. 9, no. 8, Aug. 2023, Art. no. 889, https://doi.org/10.3390/horticulturae9080889.

[11] M. Mohd Ali, N. Hashim, S. Abd Aziz, and O. Lasekan, "Characterisation of Pineapple Cultivars under Different Storage Conditions Using Infrared Thermal Imaging Coupled with Machine Learning Algorithms," *Agriculture*, vol. 12, no. 7, Jul. 2022, Art. no. 1013, https://doi.org/10.3390/agriculture12071013.

[12] N. C. Kundur and P. B. Mallikarjuna, "Deep Convolutional Neural Network Architecture for Plant Seedling Classification," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9464–9470, Dec. 2022, https://doi.org/10.48084/etasr.5282.

[13] W. N. S. Rahimi, M. A. H, and M. S. A. M. Ali, "Ananas comosus crown image thresholding and crop counting using a colour space transformation scheme," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 18, no. 5, pp. 2472–2479, Oct. 2020, https://doi.org/10.12928/telkomnika.v18i5.13895.

[14] N. H. H. Cuong, T. H. Trinh, P. Meesad, and T. T. Nguyen, "Improved YOLO object detection algorithm to detect ripe pineapple phase," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 1, pp. 1365–1381, Jan. 2022, https://doi.org/10.3233/JIFS-213251.

[15] M. A. A. Nawawi, F. S. Ismail, and H. Selamat, "Comprehensive Pineapple Segmentation Techniques with Intelligent Convolutional Neural Network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 3, pp. 1098–1105, Jun. 2018, https://doi.org/10.11591/ijeecs.v10.i3.pp1098-1105.

[16] J. Hobbs, P. Prakash, R. Paull, H. Hovhannisyan, B. Markowicz, and G. Rose, "Large-Scale Counting and Localization of Pineapple Inflorescence Through Deep Density-Estimation," *Frontiers in Plant Science*, vol. 11, Jan. 2021, Art. no. 599705, https://doi.org/10.3389/fpls.2020.599705.

[17] A. N. J. Kukunuri and D. Singh, "Efficient Application of Drone with Satellite data for Early-Stage Wheat Detection: For Precision Agriculture Monitoring," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Kuala Lumpur, Malaysia, 2022, pp. 4388–4391, https://doi.org/10.1109/IGARSS46834.2022.9883266.

[18] A. Agarwal, A. K. Singh, S. Kumar, and D. Singh, "Critical analysis of classification techniques for precision agriculture monitoring using satellite and drone," in *2018 IEEE 13th International Conference on Industrial and Information Systems*, Rupnagar, India, 2018, pp. 83–88, https://doi.org/10.1109/ICIINFS.2018.8721422.

[19] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data*, vol. 11, no. 1, Jan. 2024, Art. no. 5, https://doi.org/10.1186/s40537-023-00863-9.

[20] Z. Li, Y.-W. Chong, M. N. Ab Wahab, G.-K. Lim, and R. Dawood, "Classification and Prediction of Pineapple Quality using Deep Learning," in *2023 4th International Conference on Big Data Analytics and Practices*, Bangkok, Thailand, 2023, pp. 1–6, https://doi.org/10.1109/IBDAP58581.2023.10271948.

[21] C. Wang, J. Zhou, C. Xu, and X. Bai, "A Deep Object Detection Method for Pineapple Fruit and Flower Recognition in Cluttered Background," in *2nd International Conference on Pattern Recognition and Artificial Intelligence*, Zhongshan, China, 2020, pp. 218–227, https://doi.org/10.1007/978-3-030-59830-3_19.

[22] C.-Y. Chang, C.-S. Kuan, H.-Y. Tseng, P.-H. Lee, S.-H. Tsai, and S.-J. Chen, "Using deep learning to identify maturity and 3D distance in pineapple fields," *Scientific Reports*, vol. 12, no. 1, May 2022, Art. no. 8749, https://doi.org/10.1038/s41598-022-12096-6.

[23] D. Hussain, I. Hussain, M. Ismail, A. Alabrah, S. S. Ullah, and H. M. Alaghbari, "A Simple and Efficient Deep Learning-Based Framework for Automatic Fruit Recognition," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, Feb. 2022, Art. no. 6538117, https://doi.org/10.1155/2022/6538117.

[24] Y. Chen, S. Pan, and H. Wang, "Performance comparison of different convolutional neural networks for vegetable and fruit recognition," *Applied and Computational Engineering*, vol. 5, pp. 593–602, May 2023, https://doi.org/10.54254/2755-2721/5/20230652.

[25] Ma. C. V. Magabilin, A. C. Fajardo, and R. P. Medina, "Optimal Ripeness Classification of the Philippine Guyabano Fruit using Deep Learning," in *2022 Second International Conference on Power, Control and Computing Technologies*, Raipur, India, 2022, pp. 1–5, https://doi.org/10.1109/ICPC2T53885.2022.9777014.

[26] A. Kaur, V. Kukreja, P. Tiwari, M. Manwal, and R. Sharma, "Fruitful Fusion: An Accuracy-Boosting Ensemble of VGG19 and Convolutional Neural Networks for Dragon Fruit Classification," in *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation*, Gwalior, India, 2024, vol. 2, pp. 1–5, https://doi.org/10.1109/IATMSI60426.2024.10502785.

[27] D. Mohapatra, N. Das, K. K. Mohanty, and J. Shresth, "Automated Visual Inspecting System for Fruit Quality Estimation Using Deep Learning," in *2nd International Conference on Innovation in Electrical Power Engineering, Communication, and Computing Technology*, Bhubaneswar, India, 2021, pp. 379–389, https://doi.org/10.1007/978-981-16-7076-3_33.

[28] A. Bandyopadhyay, S. Ghosh, M. Bose, L. Kessi, and L. Gaur, "Supervised Neural Networks for Fruit Identification," in *5th International Conference on Recent Trends in Image Processing and Pattern Recognition*, Kingsville, TX, USA, 2022, pp. 220–230, https://doi.org/10.1007/978-3-031-23599-3_16.

[29] M. S. Patil and P. B. Mane, "Fused Image Classification using Pre-trained Deep Convolutional Neural Networks," in *2023 Third International Conference on Artificial Intelligence and Smart Energy*, Coimbatore, India, 2023, pp. 1215–1221, https://doi.org/10.1109/ICAIS56108.2023.10073743.

[30] D. Hindarto, N. Afarini, and E. T. E. H, "Comparison Efficacy of VGG16 and VGG19 Insect Classification Models," *JIKO (Jurnal Informatika dan Komputer)*, vol. 6, no. 3, pp. 189–195, Dec. 2023, https://doi.org/10.33387/jiko.v6i3.7008.

[31] M. Widyaningsih, T. K. Priyambodo, M. E. Wibowo, and M. Kamal, "Optimization Contrast Enhancement and Noise Reduction for Semantic Segmentation of Oil Palm Aerial Imagery," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 1, pp. 597–609, Feb. 2023, https://doi.org/10.22266/ijies2023.0228.51.

[32] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," *AgriEngineering*, vol. 4, no. 4, pp. 871–887, Dec. 2022, https://doi.org/10.3390/agriengineering4040056.

[33] R. Wan Nurazwin Syazwani, H. Muhammad Asraf, M. A. Megat Syahirul Amin, and K. A. Nur Dalila, "Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning," *Alexandria Engineering Journal*, vol. 61, no. 2, pp. 1265–1276, Feb. 2022, https://doi.org/10.1016/j.aej.2021.06.053.